



doi:10.7659/j.issn.1005-6947.2020.02.009
http://dx.doi.org/10.7659/j.issn.1005-6947.2020.02.009
Chinese Journal of General Surgery, 2020, 29(2):179-189.

· 基础研究 ·

基于免疫相关基因标签的原发性肝癌预后评分系统的建立

彭颖¹, 龚光辉¹, 李景和^{1,2}, 王俊普^{1,2}

(1. 中南大学湘雅医院 病理科, 湖南 长沙 410008; 2. 中南大学基础医学院 病理学系, 湖南 长沙 410013)

摘要

背景与目的: 原发性肝癌 (HCC) 是最常见的预后较差的恶性肿瘤之一, 其病因与发病机制尚未完全明了, 因此, 寻找 HCC 患者可靠的预后指标与生存生物标志物具有重要的临床价值。本研究通过生物信息学方法筛选 HCC 预后免疫相关基因, 并构建基于免疫相关基因标签的预后风险评分模型, 为 HCC 患者的预后评估及个体化治疗的临床决策提供依据。

方法: 从 TCGA 数据库获取 HCC 患者的临床信息以及 RNA-seq 数据 (377 个 HCC 样本和 50 个相邻的非癌组织样本)。在 Immport 数据库中下载免疫相关基因资料, 使用 R 语言的 limma 包在 HCC 组织中筛选出差异常表达的免疫相关基因。利用单变量和多变量 Cox 比例风险回归模型鉴定出与 HCC 患者 (377 个 HCC 患者中临床资料完整的 344 例) 总体生存率 (OS) 密切相关的免疫相关基因, 并以此构建基于基因标签的预后风险评分模型, 对 HCC 患者预后风险进行评分。同时从上述模型样本中随机抽取 50% 的病例 (172 例) 为验证样本行内部验证。用 Kaplan-Meier 方法分析高风险分与低风险分患者的生存状态, 用 ROC 曲线以及 C-index 分析评估该风险评分的准确性。最后, 分析该风险评分与 HCC 临床病理特征的关系, 并采用单因素和多因素 Cox 回归分析, 确定该风险评分作为独立预后因素的有效性。

结果: 在 HCC 的癌和癌旁组织中鉴定出 329 个差异表达的免疫相关基因, 其中 24 个与 HCC 的 OS 有关 (均 $P < 0.001$)。使用前向和后向选择算法进行了多变量 Cox 比例风险回归分析确定 PSMD14、S100A11、FABP6、RBP2、LCNL1、FCN2、NDRG1、CSPG5 和 NR6A1 为 OS 的 9 个高风险基因。按此 9 基因标签预后风险评分模型划分, 模型样本中高风险分 HCC 患者 OS 明显差于低风险分 HCC 患者 ($P = 1.715E-08$), 内部验证样本中得到验证结果相同 ($P = 2.222E-05$)。模型样本风险评分模型的 ROC 下曲线面积 (AUC) 在 1、3 年时分别为 0.790 和 0.733, 内部验证样本分别为 0.799 和 0.743, C-index 分析结果显示, 模型样本以及验证样本的 C-index 分别为 0.715 (95% CI=0.683~0.829) 和 0.756 (95% CI= 0.668~0.762)。HCC 的肿瘤分级, 病理分期, T 分期和新肿瘤事件的发生与风险评分明显有关 (均 $P < 0.05$)。单因素和多因素 Cox 分析显示, 风险评分是 HCC 的独立预后因素 (单因素: $HR = 1.057$, 95% CI=1.041~1.074, $P < 0.001$; 多因素: $HR = 1.050$, 95% CI=1.033~1.067, $P < 0.001$)。

结论: 通过 TCGA 数据库挖掘, 鉴定出 9 个与 HCC 患者预后密切相关的免疫相关基因, 且构建了基于 9 基因标签的预后风险评分系统, 该系统有助于临床医生判断 HCC 患者预后和制定个性化的治疗方案。

关键词

癌, 肝细胞; 生物标记, 肿瘤; 免疫基因; 预后; 计算生物学

中图分类号: R735.7

基金项目: 国家自然科学基金资助项目 (81602167); 湖南省自然科学基金资助项目 (2017JJ3494; 13JJ6020)。

收稿日期: 2019-10-28; **修订日期:** 2020-01-16。

作者简介: 彭颖, 中南大学湘雅医院住院医师, 主要从事消化系统肿瘤方面的研究。

通信作者: 王俊普, Email: wang-jp2013@csu.edu.cn

Establishment of a prognostic risk score model of hepatocellular carcinoma based on an immune-related gene signature

PENG Ying¹, GONG Guanghui¹, LI Jinghe^{1,2}, WANG Junpu^{1,2}

(1. Department of Pathology, Xiangya Hospital, Central South University, Changsha 410008, China; 2. Department of Pathology, School of Basic Medicine, Central South University, Changsha, 410013, China)

Abstract

Background and Aims: Hepatocellular carcinoma (HCC) is one of the most common malignant tumors with poor prognosis, and its etiology and pathogenesis are still elusive. Therefore, identification of reliable prognostic factors and survival biomarkers of HCC patients is of great clinical importance. This study was to screen the prognostic immune-related genes of HCC through bioinformatics approach, and then construct a prognostic risk score model based on an immune-related gene signature, so as to provide a basis for prognosis evaluation and individualized treatment decision-making regarding the HCC patients.

Methods: The clinical information and RNA-seq data (377 HCC samples and 50 adjacent non-cancerous samples) of HCC patients were obtained from the TCGA database. The information of immune-related genes was downloaded from the Immport database, and the differentially expressed immune-related genes were selected from HCC tissues using the limma package of R software. The immune-related genes closely related to the overall survival (OS) among the HCC patients (344 cases with complete clinical record in the 377 HCC patients) were determined by univariate and multivariate Cox proportional risk regression models, and then the prognostic risk score model based on an immune-related gene signature was constructed by using the determined genes, by which the prognostic risks of the HCC patients were scored. Meanwhile, 50% cases (172 cases) were randomly picked up from the above model sample as an internal validation sample for internal validation. Kaplan-Meier method was used to analyze the survival status between patients with high and low risk score, and the accuracy of the risk score was evaluated by ROC curve and C-index analysis. Finally, the relations of the risk score with the clinicopathologic factors of HCC were analyzed, and the effectiveness of this risk score as an independent prognostic risk factor for HCC was ascertained by univariate and multivariate Cox regression analysis.

Results: A total of 329 differentially expressed immune-related genes between HCC tissue and tumor adjacent were identified, in which 24 were significantly associated with the OS of the HCC patients (all $P < 0.001$), and 9 genes that included PSMD14, S100A11, FABP6, RBP2, LCNL1, FCN2, NDRG1, CSPG5 and NR6A1 were determined as high-risk genes for OS by using forward and backward selection algorithm in multivariate Cox proportional hazards regression analysis. According to the classification by the 9-gene signature based prognostic risk score model, the OS in patients with high risk score was significantly worse than that in patients with low risk score in the model sample ($P = 1.715 \times 10^{-8}$), and the same result was also obtained in the internal validation sample ($P = 2.222 \times 10^{-5}$). The 1- and 3-year values of the area under the ROC curve (AUC) were 0.790 and 0.733 in the model sample, and were 0.799 and 0.743 in the internal validation sample, respectively. The results of C-index analysis showed that the C-index values in the model sample and the internal validation sample were 0.715 (95% CI=0.683–0.829) and 0.756 (95% CI=0.668–0.762), respectively. Tumor grade, pathological stage, T stage and new tumor events were correlated with the risk score (all $P < 0.05$). Univariate and multivariate Cox analysis showed that the risk score was an independent prognostic factor for HCC (univariate: HR=1.057, 95% CI=1.041-1.074, $P < 0.001$; multivariate: HR=1.050, 95% CI=1.033-1.067, $P < 0.001$).

Conclusion: Nine immune-related genes closely related to the prognosis of HCC patients are identified by TCGA database mining, and a 9-gene signature based prognostic risk score model is developed, which may help the clinicians to assess the prognosis of the HCC patients and design a personalized treatment plan for them.

Key words

Carcinoma, Hepatocellular; Biomarkers, Tumor; Immune Genes; Prognosis; Computational Biology

CLC number: R735.7

肝细胞癌 (hepatocellular carcinoma, HCC) 是全球第六大最常见的恶性肿瘤,也是癌症相关死亡的第四大诱因^[1]。HCC的机理与基因突变的增量积累相关,从而产生异常的免疫应答、细胞周期、和血管生成^[2]。至于免疫应答,效应免疫细胞可以对HCC进行免疫控制,从而有效减少恶性转化细胞。然而,HCC进展过程清楚地显示肿瘤免疫控制的失败会抑制抗癌免疫反应^[3]。

随着医疗技术的迅速发展,HCC的诊断和治疗已大大改善,但预后仍然很差,5年生存率低于40%。尽管有多种预后生物标志物,包括甲胎蛋白(AFP)、DES- γ 羧基凝血酶、和细胞角蛋白19被广泛应用于临床实践中^[4],但它们的有效性会被各种因素影响,而且单个指标的预测能力不足。相反,基因标签提供更好的预测性能,多基因预后标签可以指导临床医师选择适当的治疗^[5]。

最近,能够促进免疫系统攻击癌细胞的免疫检查点抑制剂的开发取得了显著成功^[6],在PubMed数据库和中国知网中检索肝细胞癌、免疫基因、预后模型等关键词,尚未有研究构建HCC的免疫相关基因标签预后模型。

因此本研究基于The Cancer Genome Atlas (TCGA) 数据库,对于HCC鉴定有效且可靠的免疫基因相关生物标志物以进行预后预测。利用R语言在HCC中挖掘差异表达的免疫基因,结合TCGA数据库中HCC临床样本并分析患者的预后,选择其中与预后明显相关的免疫基因构建可以预测HCC预后的基因标签。鉴定模型中的核心免疫基因和途径可以增强临床应用,提供新的治疗靶点并产生有关癌症进展的新见解。

1 材料与方法

1.1 数据来源及研究内容

完整的mRNA表达数据和临床信息可从TCGA数据库 (<https://cancergenome.nih.gov>) 获得。本次研究包括377个HCC样本和50个相邻的非癌性组织样本。提取相应的临床信息,包括年龄、性别、TNM分类、病理分期、组织学分级、肿瘤状态等。从Immport数据库 (<https://www.immport.org/home>) 下载免疫基因相关信息。使用R语言的

limma包提取差异表达的免疫相关差异基因。除临床信息不完整,随访信息<30 d的患者,用单变量Cox回归分析评估与344例HCC患者总体生存率(OS)显著相关的免疫基因,然后使用多元Cox比例风险回归分析进行进一步筛选,计算每个基因的危险比(HR)。然后,根据表达水平的线性组合构建预后风险评分模型,该风险评分由从多元Cox回归分析获得的回归系数加权,使用中位风险评分将344例患者分为高风险分組和低风险分組。同时从上述344例模型样本中随机抽取50%的病例(172例)为验证样本行内部验证。使用Kaplan-Meier曲线比较两组的预后,并使用对数秩检验评估差异。绘制ROC曲线以分析预测患者总体生存的有效性。用Z检验检查ROC曲线的曲线下面积(AUC)和一致性指数(C-index)分析模型的准确性。进行单因素和多因素Cox回归分析以研究特征性风险评分与临床病理特征之间的关系。

1.2 统计学处理

以上所有统计分析均使用R 3.6.1 (www.r-project.org) 进行, $P<0.05$ 被认为具有统计学意义。

2 结果

2.1 HCC中差异表达的免疫基因鉴定

使用Wilcoxon秩和检验分析了377个HCC和50个正常肝脏样品中2 498个免疫相关基因的表达,并鉴定了329个差异表达的免疫相关基因,包括267个上调的免疫相关基因和62个下调的免疫相关基因($FDR<0.05$, $|\log_2FC|>1$) (图1)。

2.2 与HCC生存相关的免疫相关基因鉴定

对HCC中329个差异表达的免疫相关基因的表达进行了单变量Cox回归分析,以鉴定预后性差异免疫相关基因。数据显示,24个差异表达的免疫相关基因的表达与HCC患者的OS明显有关(均 $P<0.001$)。为了确定预测预后的最佳模型,使用前向和后向选择算法进行了多变量Cox比例风险回归分析。PSMD14、S100A11、FABP6、RBP2、LCNL1、FCN2、NDRG1、CSPG5和NR6A1被确定为OS模型中的风险基因,高风险基因与患者的预后成负相关。模型中的基因全为高风险基因(表1)。

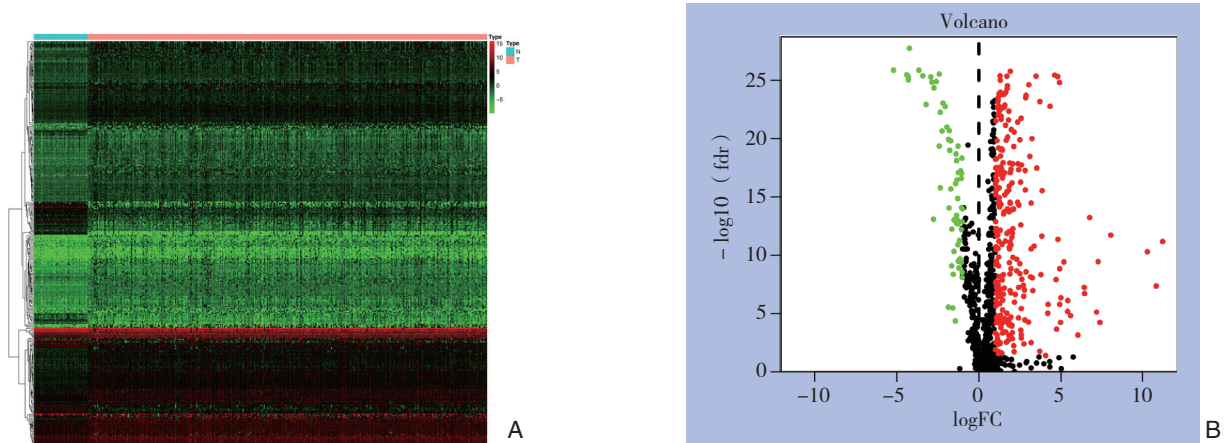


图 1 差异表达免疫相关基因分析 A: 免疫相关基因在 HCC 组织和癌旁组织的差异表达热图; B: 火山图显示 HCC 组织样品 ($n=377$) 中 2 498 个免疫相关基因与正常肝脏样品 ($n=50$) 的差异表达 (红点代表 267 个上调的基因, 绿点代表 62 个下调的基因, 其余黑点表示无差异表达的基因)

Figure 1 Analysis of the differentially expressed immune-related genes A: Heat map of differentially expressed genes between HCC tissue and tumor adjacent tissue; B: Volcano showing the differential expressions of 2 498 immune-related genes in HCC tissue samples ($n=377$) compared with normal liver tissue sample ($n=50$) (the red dots standing for the 267 up-regulated genes, the green dot standing for the down-regulated genes and the black dots standing for the rest genes without differential expression)

表 1 HCC 中 9 个 OS 高风险免疫相关基因

Table 1 The 9 high-risk immune-related genes associated with the OS of HCC

基因	影响系数	HR (95% CI)	P
PSMD14	0.076	1.079 (1.012~1.150)	2.011E-02
S100A11	0.001	1.001 (1.000~1.002)	8.005E-02
FABP6	0.058	1.060 (0.995~1.129)	7.129E-02
RBP2	0.019	1.019 (1.008~1.031)	8.154E-04
LCNL1	0.161	1.175 (0.997~1.386)	5.491E-02
FCN2	0.028	1.028 (1.019~1.037)	1.103E-10
NDRG1	0.006	1.006 (1.003~1.010)	7.054E-04
CSPG5	0.450	1.569 (1.148~2.145)	4.728E-03
NR6A1	0.157	1.170 (1.012~1.353)	3.399E-02

2.3 基于免疫基因标签的预后风险评分模型的建立与验证

选择上述 9 个 OS 相关的免疫相关基因构成基因标签。根据以下公式计算每个患者的风险评分: 风险评分 = $(0.076) \times \text{PSMD14 值} + (0.001) \times \text{S100A11 值} + (0.058) \times \text{FABP6 值} + (0.019) \times \text{RBP2 值} + (0.161) \times \text{LCNL1 值} + (0.028) \times \text{FCN2 值} + (0.006) \times \text{NDRG1 值} + (0.450) \times \text{CSPG5 值} + (0.157) \times \text{NR6A1 值}$ 。根据风险评分对患者进行分类, 将其分为高风险分部和低风险分部。患者的风险评分分布和生存状况如图 2A-B 所示。高风险分部的病死率为 48.8%, 而低风险分部的病死率为 26.6%。两组之间的病死率存在明

显差异 ($P=2.722E-07$)。Kaplan-Meier 曲线显示, 高风险分部的患者预后较差 (对数秩检验: $P=1.715E-08$) (图 2C)。高风险分部的中位生存期为 2.753 年, 高危患者的 3 年和 5 年 OS 率分别为 48.8% 和 33.8%, 低风险分部的中位生存期为 6.937 年, 而低危患者的 3、5 年 OS 率分别为 76.8% 和 62.7%。使用随时间变化的 ROC 曲线, 测量了 1、3 年 OS 的预后风险模型的预测性能。模型样本风险评分在 ROC 的 AUC 值下的面积在 1、3 年时分别为 0.790、0.733 (图 2D-E)。

对该模型进行内部验证, 利用 R 语言将 344 个 HCC 样本中, 随机抽取 172 个样本, 根据该模型的风险评分对内部验证样本患者进行分类, 将其分为高风险分部和低风险分部。内部验证样本的风险评分分布和生存状况如图 3A-B 所示。内部验证样本中, 高风险分部的病死率为 46.5%, 而低风险分部的病死率为 19.8%。两组之间的病死率存在明显差异 ($P=1.951E-04$)。Kaplan-Meier 曲线显示, 高风险分部患者预后明显差于低风险分部患者 (对数秩检验: $P=2.22E-05$) (图 3C)。高风险分部的中位生存期为 1.284 年, 高危患者的 3、5 年 OS 率分别为 48.8%、36.4%, 低风险分部的中位生存期超过 10 年, 而低危患者的 3、5 年 OS 率分别为 77.4%、62.8%。内部验证样本运用风险评分模型得到一致的生存分析结果。内部验证样本的风险

评分在ROC的AUC值下的面积在1、3年时分别为0.799和0.743(图3D-E),均证明此模型具有较高的准确性。

C-index是在ROC曲线下面积的基础上发展而来,用于估计预测结果与实际观察结果相一致

的概率。在R语言中用C-index分析此模型的准确性,在模型样本和内部验证样本中,C-index的值分别为0.715(95%CI=0.668~0.762)和0.756(95%CI=0.683~0.829),故此风险评分模型具有较高的准确性。

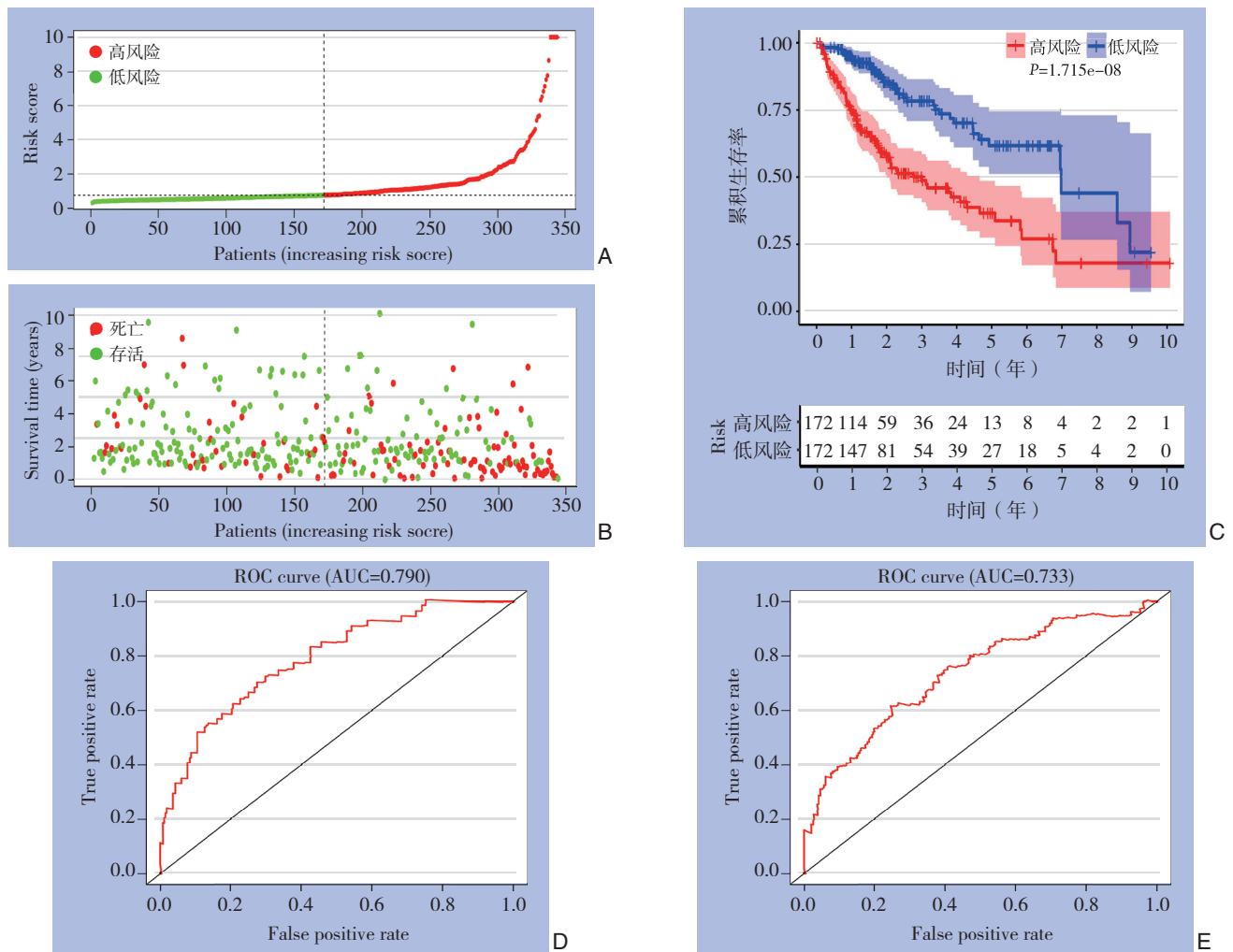


图2 模型样本高风险分和低风险分HCC患者的分析 A: 高风险分(红色)和低风险分(绿色)患者的风险评分分布; B: HCC患者的生存状况(红点表示死亡的患者,绿点表示存活的患者); C: 高风险分(红线)和低风险分(蓝线)患者的OS的Kaplan-Meier生存曲线; D: 时间依赖性ROC曲线显示HCC患者在1年OS时的AUC值; E: 时间依赖性ROC曲线显示HCC患者在3年OS时的AUC值

Figure 2 Analysis of HCC patients with high-risk score and low-risk score in the model sample A: Distribution of the patients with high-risk score (red color) and low-risk score (green color); B: Survival status of the HCC patients (the red dots standing for the dead cases and the green dots standing for the survivors); C: The OS Kaplan-Meier curves of patients with high-risk score (red line) and low-risk score (blue line); D: Time-dependent ROC curve showing the AUC of 1-year OS of the HCC patients; E: Time-dependent ROC curve showing the AUC of 3-year OS of the HCC patients

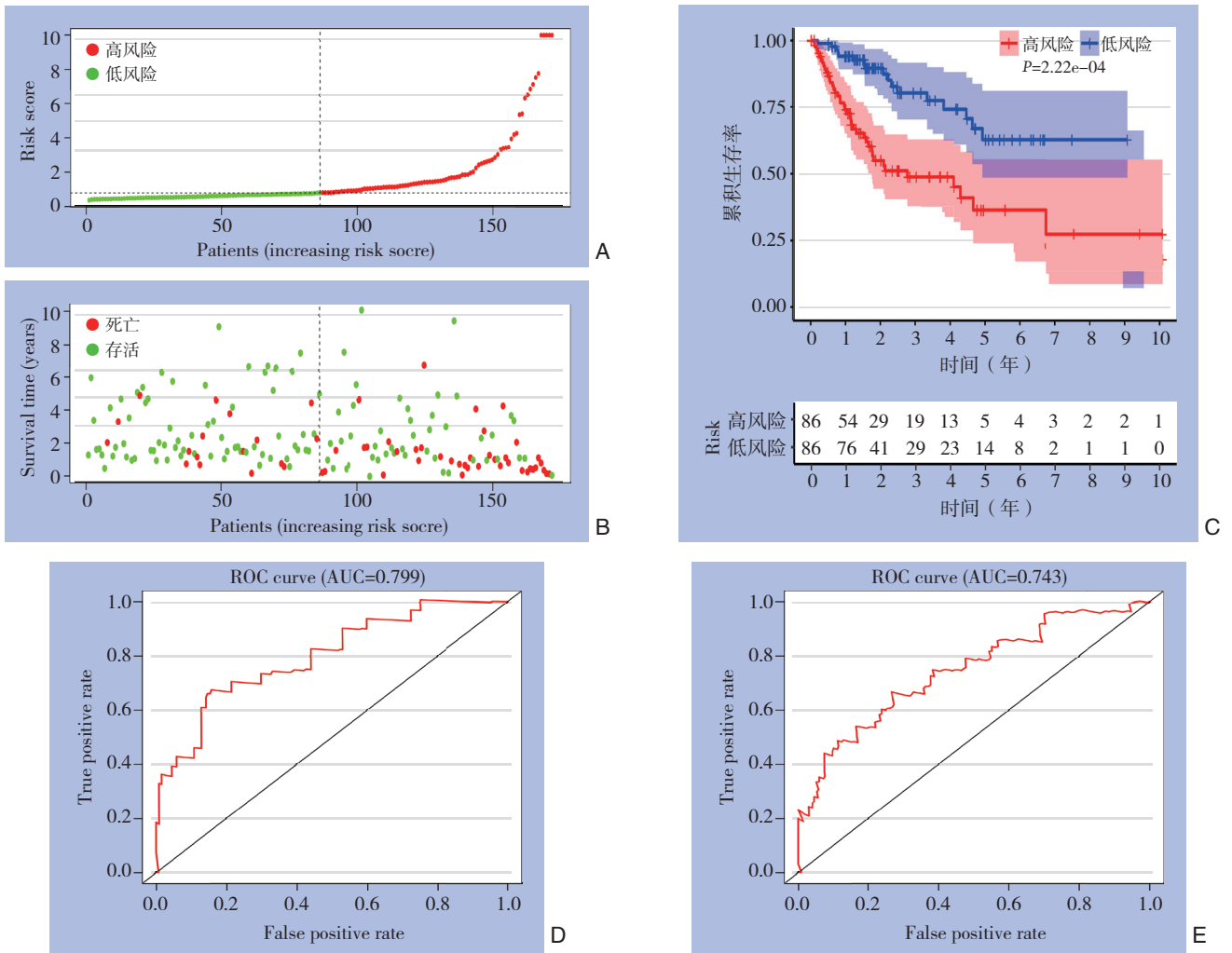


图 3 内部验证样本中高风险和低风险 HCC 患者的分析 A: 高风险分(红色)和低风险分(绿色)患者的风险评分分布; B: HCC 患者的生存状况(红点表示死亡的患者, 绿点表示存活的患者); C: 高风险分(红线)和低风险分(蓝线)患者的 OS 的 Kaplan-Meier 生存曲线; D: 时间依赖性 ROC 曲线显示 HCC 患者在 1 年 OS 时的 AUC 值; E: 时间依赖性 ROC 曲线显示 HCC 患者在 3 年 OS 时的 AUC 值

Figure 3 Analysis of HCC patients with high-risk score and low-risk score in the internal validation sample A: Distribution of the patients with high-risk score (red color) and low-risk score (green color); B: Survival status of the HCC patients (the red dots standing for the dead cases and the green dots standing for the survivors); C: The OS Kaplan-Meier curves of patients with high-risk score (red line) and low-risk score (blue line); D: Time-dependent ROC curve showing the AUC of 1-year OS of the HCC patients; E: Time-dependent ROC curve showing the AUC of 3-year OS of the HCC patients

2.4 基于免疫基因标签的预后风险评分与临床病理特征的关系

进一步研究风险评分分布与临床参数分层之间的关系, 肿瘤分级、病理分期、T 分期和新肿瘤事件与风险评分有关(均 $P < 0.05$) (图 4); 其他临床特征, 例如年龄、性别和肿瘤状态对风险评分没有影响(均 $P > 0.05$)。使用单变量和多变量 Cox 回归来检查风险评分与其他临床病理特征

的组合能力, 以预测生存率。单因素 Cox 回归显示, 肿瘤的 T 分期、病理分期和特征性风险评分与生存率有关(均 $P < 0.05$); 而多因素回归分析显示, 只有风险评分与预后明显有关, 单因素与多因素分析均有统计学意义(单因素: $HR = 1.057$, $95\% CI = 1.041 \sim 1.074$, $P < 0.001$; 多因素: $HR = 1.050$, $95\% CI = 1.033 \sim 1.067$, $P < 0.001$), 这意味着风险评分是 HCC 的独立预后指标(表 2)。

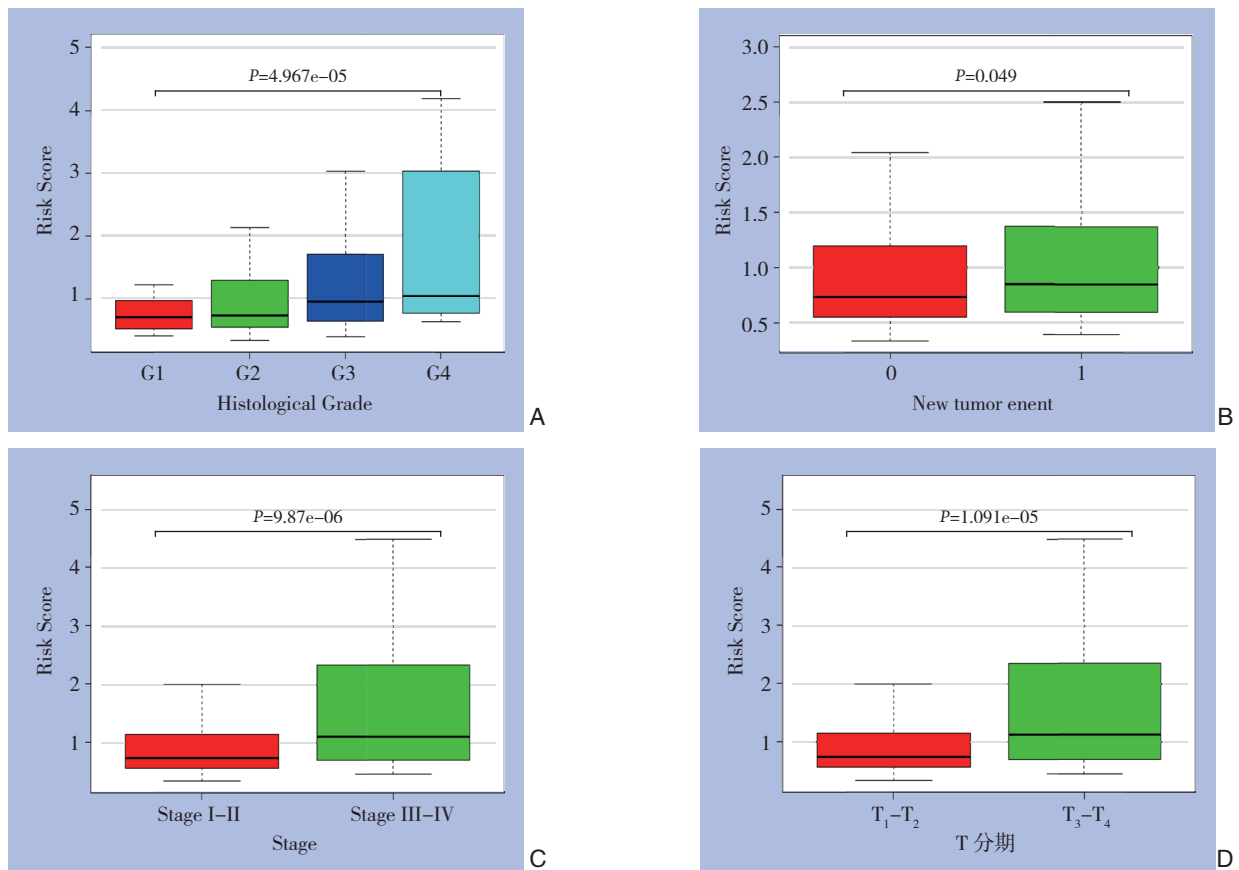


图 4 风险评分分布与临床参数之间的关系 A: 组织学分级; B: 新肿瘤事件; C: 病理学分期; D: T 分期

Figure 4 Relationship between risk score distribution and clinical factors A: Histological grade; B: New tumor events; C: Pathologic stage; D: T stage

表 2 HCC 患者 OS 的单因素和多因素 Cox 回归分析。

Table 2 Univariate and multivariate Cox regression analysis of OS in HCC patients

临床变量	单因素分析		多因素分析	
	HR (95% CI)	P	HR (95% CI)	P
年龄	1.005 (0.991~1.020)	4.745E-01	1.005 (0.990~1.020)	5.352E-01
性别	0.754 (0.511~1.113)	1.558E-01	0.867 (0.581~1.293)	4.830E-01
病理学分期	1.806 (1.461~2.232)	4.607E-08	1.074 (0.439~2.627)	8.758E-01
T 分期	1.778 (1.454~2.174)	2.123E-08	1.619 (0.690~3.797)	2.680E-01
组织学分级	1.118 (0.862~1.450)	4.005E-01	1.169 (0.887~1.541)	2.669E-01
风险评分	1.057 (1.041~1.074)	8.002E-12	1.050 (1.033~1.067)	7.882E-09

3 讨 论

HCC 是高度异质性疾病^[7]。数据显示, HCC 的发病率和病死率在世界范围内一直在增加。尽管在治疗上已取得很大进展, 但 HCC 的总体生存率仍然很差。因此, 迫切需要鉴定肝癌的敏感和特异性生物标志物。有证据^[8-9]表明, 常用的临床病理学参数 (例如 TNM 分期, 年龄, 性别, 病毒感染和血清 AFP 水平) 不足以准确预测患者的预后。因此, 在癌症的发展和预后过程中, 大量的 mRNA

被检查并鉴定为生物标志物。例如, Gao 等^[10]研究 Rho 鸟嘌呤核苷酸交换因子 39 (ARHGEF39) 的表达及其在 HCC 中的预后意义。Gu 等^[11]发现 CELSR3 mRNA 表达在 HCC 中增加, 提示预后不良。但是, 由于这些生物标志物会受到许多因素的影响, 预测能力不稳定。

大多数 HCC 发生在慢性炎症的背景下, 许多患者还伴有肝硬化, 局部和全身免疫缺陷。此外, 肝脏本身就是免疫器官, 肝脏独特的免疫生物学特性可以促进肿瘤的发生和对肿瘤的耐受性^[12],

同时也为免疫检查点抑制剂的靶向治疗提供了机会^[13]。

多年来,随着高通量测序技术的发展,微阵列数据和生物信息学的结合被广泛用于多种癌症的诊断以及预后的生物标志物的开发。利用数据挖掘技术以及单变量和多变量Cox回归分析来开发包含各种相关基因的基因标签。这样的基因标签被广泛用于分子诊断,个体化治疗,准确的存活预测^[14],其预测准确性也优于单一生物标志物^[15]。

因此利用生物信息学技术,建立一个与免疫相关的基因标签以指导HCC患者的治疗和预后十分必要。本研究利用Immport数据库,筛选出在HCC癌和癌旁差异表达的329个IRG,通过单变量和多变量Cox比例风险回归模型来鉴定出了9个具有OS的预后价值的免疫相关基因(PSMD14、S100A11、FABP6、RBP2、LCNL1、FCN2、NDRG1、CSPG5和NR6A1),这9个基因在HCC的肿瘤组织中都是被上调,且与不良预后成正相关。利用COX回归模型得出基因标签的风险评分。根据风险评分,将患者分为高风险分組和低风险分組。低风险分組患者的预后明显好于高风险分組。ROC曲线、C-index分析以及内部验证表明此风险评分系统在预测HCC总生存期时具有较准确的预测性能。此外,将临床病理因素与风险评分进行相关性分析,结果表明,风险评分与患者T分期,组织学分级,以及病理学分期和是否发生新肿瘤事件等成负相关。最后,本研究进行了单因素和多因素Cox回归分析,以探索风险评分和其他临床病理参数共同预测生存的能力。结果证实风险评分是一个稳定,独立的预后指标。

在本研究鉴定的免疫基因标签中,去泛素化酶PSMD14,也称为RPN11、POH1和PAD1,是26S蛋白酶体中19S调节帽的组成部分。已证明其参与多种生物学过程的调控,包括细胞生存力^[16],衰老^[17],多能性^[18],细胞自噬^[19]和DNA损伤^[20]。去泛素化酶PSMD14通过去泛素化和稳定化E2F1,促进了HCC的发生^[21]。Wang等^[22]发现,PSMD14可促进HCC细胞中TGF- β 信号传导增加肝癌细胞的转移潜能与不良预后相关。S100A11是包含2个EF手钙结合基序的S100蛋白家族的成员。该蛋白可能在运动,侵袭和微管蛋白聚合中起作用^[23]。S100A11是伪足突出以及肿瘤细胞迁移和侵袭的必需蛋白之一。据报道^[25],S100A11是胃

癌淋巴结转移的准确预测因子^[24]和结肠源性肝转移的预测因子。TGF- β 可诱导肝癌中S100A11的表达^[26]。Luo等^[27]发现S100A11在肝癌中高表达,并在EGFRvIII-STAT3途径中起着至关重要的作用,以促进癌细胞的侵袭和转移。RBP2属于JARID家族,可以使H3K4me2和H3K4me3显著脱甲基^[28],RBP2的功能主要与癌症的上皮-间质转化(EMT)、迁移、侵袭和细胞增殖有关。RBP2可能通过PI3K / AKT / HIF-1 α 信号传导参与促进VEGF表达,RBP2与VEGF协同参与了HCC肿瘤血管生成^[29]。FCN2基因编码无花果酶2(L-ficolin,FCN2),是一种可溶性血清蛋白,被认为在宿主先天免疫中发挥重要作用,主要在肝脏中表达,并分泌入血液循环^[30]。此外,据报道FCN2在TGF- β 信号通路在调节中扮演至关重要的角色,从而参与乳腺癌、胰腺癌、结肠癌和HCC等^[31-32]发生进展和转移。NDRG1是N-Myc下游调控的基因1,在HCC中,NDRG1通过直接与GSK-3 β 和Nur77相互作用来阻止 β -catenin降解而成为癌基因^[33]。Shi等^[34]发现NDRG1通过抑制 β -catenin的降解促进核易位和增强的CSC样特性而促进了HCC转移的潜力,NDRG1的上调与HCC的预后不良有关。

尚未有研究证明FABP6(回肠脂质结合蛋白;回肠胆汁酸结合蛋白)、LCNL1(脂质蛋白样1)和NR6A1(孤儿受体)在肝癌中发生发展中的关系。FABP6参与结肠上皮细胞内的胆汁酸胞内运输,Ohmachi等^[35]发现了FABP6在结直肠癌中过表达,而且Keler等^[36]发现了肝脏型脂酸性结合蛋白(L-FABP)在肝细胞内促进DNA合成从而促进细胞生长和维持细胞形态。LCNL1生物学功能涉及先天性免疫反应,并充当潜在有害亲脂分子的生理清除剂,Yang等^[37]发现LCNL1的高表达也显示HER-2阳性乳腺癌预后不良。NR6A1/CT150是一种孤儿受体,是睾丸癌(CT)抗原家族的新成员。已有研究^[38-39]表明NR6A1在改变细胞命运,包括细胞生长和分化,神经发生和生殖细胞分化等方面有重要作用。Cheng等^[40]研究表明NR6A1过表达可增强DU145和PC3细胞系的上皮-间充质转化(EMT)过程,这与前列腺癌的复发转移不良预后密切相关。FABP6、NR6A1、LCNL1在多种癌症中都起着重要作用,其HCC中的作用机制值得进一步研究。

目前尚未有研究阐明CSPG5在癌症中的作用机制,CSPG5是含鸡酸性亮氨酸的富含EGF样结构域的脑蛋白(CALEB),为硫酸软骨素蛋白聚糖(CSPG)5或神经聚糖C,是含有神经软骨素和表皮生长因子(EGF)结构域的跨膜蛋白。在神经系统的发育过程中起重要作用,Jüttner等^[41]的研究表明CSPG5参与小脑 γ -氨基丁酸能突触的突触前分化。CSPG5在其他疾病以及癌症中的作用值得进一步研究。

本研究对HCC基因表达谱进行了全面而新颖的分析,以识别可能在HCC患者的发展和预后中发挥核心作用的DEIRG。本研究的预测模型基于9个免疫相关基因的表达水平。这种方法比全基因组测序更经济和临床上可行。9基因标签结合临床病理参数,可以使临床医生更准确的分析每个患者的预后。它的风险评分系统易于理解,有助于定制治疗和医疗决策。

本研究可能为HCC的分子机制、免疫治疗和预后的预测提供了新的见解。本研究构建的新的HCC预测风险评分模型的优点之一是:预测OS方面具有很高的敏感性和特异性,并且随机内部验证也证明其有效性。此外,风险评分模型与免疫抑制环境和免疫检查点表达相关,从而帮助临床医生为HCC患者选择个性化免疫疗法。

同时,本研究也存在一些局限性。首先,风险评分模型需要在多中心临床试验和前瞻性研究中得到进一步验证。其次,仍需进一步对9个免疫相关基因的功能和机理进行解析。这些工作有待于进一步开展。

本文中所述的9免疫基因预后标签和基于该模型的风险评分先前尚未报道。因此,本研究使用生物信息学方法,筛选出由PSMD14、S100A11、FABP6、RBP2、LCNL1、FCN2、NDRG1、CSPG5和NR6A1组成的免疫相关基因标签,并证明其在HCC中的预后价值;该模型可以预测切除术后HCC患者的OS,可能有助于个体化治疗的临床决策。而且值得注意的是,风险评分模型提供了免疫学观点,以阐明确定HCC临床结果的机制。

参考文献

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6):394–424. doi: 10.3322/caac.21492.
- [2] Greten TF, Wang XW, Korangy F. Current concepts of immune based treatments for patients with HCC: from basic science to novel treatment approaches[J]. *Gut*, 2015, 64(5):842–848. doi: 10.1136/gutjnl-2014-307990.
- [3] Hato T, Goyal L, Greten TF, et al. Immune checkpoint blockade in hepatocellular carcinoma: current progress and future directions[J]. *Hepatology*, 2014, 60(5):1776–1782. doi: 10.1002/hep.27246.
- [4] Sun DW, Zhang YY, Sun XD, et al. Prognostic value of cytokeratin 19 in hepatocellular carcinoma: A meta-analysis[J]. *Clin Chim Acta*, 2015, 448:161–169. doi: 10.1016/j.cca.2015.06.027.
- [5] 何政, 吴慧, 郑军. 驱动蛋白家族成员15在肝癌中的表达及意义[J]. *中国普通外科杂志*, 2019, 28(1):58–63. doi:10.7659/j.issn.1005-6947.2019.01.008.
- [6] He Z, Wu H, Zheng J. Expression of kinesin family member 15 in liver cancer and its significance[J]. *Chinese Journal of General Surgery*, 2019, 28(1):58–63. doi:10.7659/j.issn.1005-6947.2019.01.008.
- [6] 陈云茹, 丹增措姆, 白文东, 等. SFPQ在肝细胞癌的表达及其对细胞增殖与细胞周期的影响[J]. *中国普通外科杂志*, 2019, 28(1):49–57. doi:10.7659/j.issn.1005-6947.2019.01.007.
- [7] Chen YR, Dan ZCM, Bai WD, et al. Expression of SFPQ in hepatocellular carcinoma and its effect on cell proliferation and cell cycle[J]. *Chinese Journal of General Surgery*, 2019, 28(1):49–57. doi:10.7659/j.issn.1005-6947.2019.01.007.
- [7] Lin DC, Mayakonda A, Dinh HQ, et al. Genomic and Epigenomic Heterogeneity of Hepatocellular Carcinoma [J]. *Cancer Res*, 2017, 77(9):2255–2265. doi: 10.1158/0008-5472.CAN-16-2822.
- [8] Pawlik TM, Poon RT, Abdalla EK, et al. Hepatitis serology predicts tumor and liver-disease characteristics but not prognosis after resection of hepatocellular carcinoma[J]. *J Gastrointest Surg*, 2004, 8(7):794–804. doi: 10.1016/j.gassur.2004.06.013.
- [9] Tan JT, Zhao C, Peng NF, et al. Association between age and overall survival of patients with hepatocellular carcinoma after hepatic resection[J]. *J Surg Oncol*, 2016, 114(8):966–970. doi: 10.1002/jso.24434.
- [10] Gao J, Jia WD. Expression of Rho Guanine Nucleotide Exchange Factor 39 (ARHGEF39) and Its Prognostic Significance in Hepatocellular Carcinoma[J]. *Med Sci Monit*, 2019, 25:7826–7835. doi: 10.12659/MSM.918270.
- [11] Gu X, Li H, Sha L, et al. CELSR3 mRNA expression is increased in hepatocellular carcinoma and indicates poor prognosis[J]. *PeerJ*, 2019, 7:e7816. doi: 10.7717/peerj.7816.
- [12] 李军良, 郭天康, 张东, 等. 转B、T淋巴细胞衰减子基因树突状细胞诱导大鼠肝移植免疫耐受的作用及机制[J]. *中国普通外科杂志*, 2019, 28(1):70–76. doi:10.7659/j.issn.1005-6947.2019.01.010.

- Li JL, Guo TK, Zhang D, et al. Immune tolerance in liver transplantation induced by dendritic cells transfected with B and T lymphocyte attenuator gene in rats and its mechanism[J]. Chinese Journal of General Surgery, 2019, 28(1):70–76. doi:10.7659/j.issn.1005-6947.2019.01.010.
- [13] Keenan BP, Fong L, Kelley RK. Immunotherapy in hepatocellular carcinoma: the complex interface between inflammation, fibrosis, and the immune response[J]. J Immunother Cancer, 2019, 7(1):267. doi: 10.1186/s40425-019-0749-z.
- [14] Yu Y, Feng X, Cang S. A two-microRNA signature as a diagnostic and prognostic marker of pancreatic adenocarcinoma[J]. Cancer Manag Res, 2018, 10:1507–1515. doi: 10.2147/CMAR.S158712.
- [15] Yuan GQ, Wei NL, Mu LY, et al. A 4-miRNAs signature predicts survival in glioblastoma multiforme patients[J]. Cancer Biomark, 2017, 20(4):443–452. doi: 10.3233/CBM-170205.
- [16] Gallery M, Blank JL, Lin Y, et al. The JAMM motif of human deubiquitinase Poh1 is essential for cell viability[J]. Mol Cancer Ther, 2007, 6(1):262–268. doi: 10.1158/1535-7163.MCT-06-0542.
- [17] Byrne A, McLaren RP, Mason P, et al. Knockdown of human deubiquitinase PSMD14 induces cell cycle arrest and senescence[J]. Exp Cell Res, 2010, 316(2):258–271. doi: 10.1016/j.yexcr.2009.08.018.
- [18] Choi J, Baek KH. Cellular functions of stem cell factors mediated by the ubiquitin-proteasome system[J]. Cell Mol Life Sci, 2018, 75(11):1947–1957. doi: 10.1007/s00018-018-2770-7.
- [19] Hao R, Nanduri P, Rao Y, et al. Proteasomes activate aggresome disassembly and clearance by producing unanchored ubiquitin chains[J]. Mol Cell, 2013, 51(6):819–828. doi: 10.1016/j.molcel.2013.08.016.
- [20] Jacquemont C, Taniguchi T. Proteasome function is required for DNA damage response and fanconi anemia pathway activation[J]. Cancer Res, 2007, 67(15):7395–7405. doi: 10.1158/0008-5472.CAN-07-1015.
- [21] Wang B, Ma A, Zhang L, et al. POH1 deubiquitylates and stabilizes E2F1 to promote tumour formation[J]. Nat Commun, 2015, 6:8704. doi: 10.1038/ncomms9704.
- [22] Wang B, Xu X, Yang Z, et al. POH1 contributes to hyperactivation of TGF-beta signaling and facilitates hepatocellular carcinoma metastasis through deubiquitinating TGF-beta receptors and caveolin-1[J]. EBioMedicine, 2019, 41:320–332. doi: 10.1016/j.ebiom.2019.01.058.
- [23] Shankar J, Messenberg A, Chan J, et al. Pseudopodial actin dynamics control epithelial-mesenchymal transition in metastatic cancer cells[J]. Cancer Res, 2010, 70(9):3780–3790. doi: 10.1158/0008-5472.CAN-09-4439.
- [24] Mori M, Shimada H, Gunji Y, et al. S100A11 gene identified by in-house cDNA microarray as an accurate predictor of lymph node metastases of gastric cancer[J]. Oncol Rep, 2004, 11(6):1287–1293.
- [25] Melle C, Ernst G, Schimmel B, et al. Colon-derived liver metastasis, colorectal carcinoma, and hepatocellular carcinoma can be discriminated by the Ca(2+)-binding proteins S100A6 and S100A11 [J]. PLoS One, 2008, 3(12):e3767. doi: 10.1371/journal.pone.0003767.
- [26] Miyazaki M, Sakaguchi M, Akiyama I, et al. Involvement of interferon regulatory factor 1 and S100C/A11 in growth inhibition by transforming growth factor beta 1 in human hepatocellular carcinoma cells[J]. Cancer Res, 2004, 64(12):4155–4161. doi: 10.1158/0008-5472.CAN-03-2750.
- [27] Luo X, Xie H, Long X, et al. EGFRvIII mediates hepatocellular carcinoma cell invasion by promoting S100 calcium binding protein A11 expression[J]. PLoS One, 2013, 8(12):e83332. doi: 10.1371/journal.pone.0083332.
- [28] Ge W, Shi L, Zhou Y, et al. Inhibition of osteogenic differentiation of human adipose-derived stromal cells by retinoblastoma binding protein 2 repression of RUNX2-activated transcription[J]. Stem Cells, 2011, 29(7):1112–1125. doi: 10.1002/stem.663.
- [29] Wang ZY, Yang J, Liu CK, et al. High Expression of Retinoblastoma-Binding Protein 2 (RBP2) in Patients with Hepatocellular Carcinoma and Its Prognostic Significance[J]. Med Sci Monit, 2017, 23:2736–2744. doi: 10.12659/msm.905262.
- [30] Le Y, Lee SH, Kon OL, et al. Human L-ficolin: plasma levels, sugar specificity, and assignment of its lectin activity to the fibrinogen-like (FBG) domain[J]. FEBS Lett, 1998, 425(2):367–370. doi: 10.1016/s0014-5793(98)00267-1.
- [31] Yang G, Liang Y, Zheng T, et al. FCN2 inhibits epithelial-mesenchymal transition-induced metastasis of hepatocellular carcinoma via TGF-beta/Smad signaling[J]. Cancer Lett, 2016, 378(2):80–86. doi: 10.1016/j.canlet.2016.05.007.
- [32] Li MO, Flavell RA. TGF-beta: a master of all T cell trades[J]. Cell, 2008, 134(3):392–404. doi: 10.1016/j.cell.2008.07.025.
- [33] Lu WJ, Chua MS, Wei W, et al. NDRG1 promotes growth of hepatocellular carcinoma cells by directly interacting with GSK-3beta and Nur77 to prevent beta-catenin degradation[J]. Oncotarget, 2015, 6(30):29847–29859. doi: 10.18632/oncotarget.4913.
- [34] Shi J, Zheng H, Yuan L. High NDRG3 expression facilitates HCC metastasis by promoting nuclear translocation of beta-catenin [J]. BMB Rep, 2019, 52(7):451–456.
- [35] Ohmachi T, Inoue H, Mimori K, et al. Fatty acid binding protein 6 is overexpressed in colorectal cancer[J]. Clin Cancer Res, 2006, 12(17):5090–5095. doi: 10.1158/1078-0432.CCR-05-2045.
- [36] Keler T, Sorof S. Growth promotion of transfected hepatoma cells by liver fatty acid binding protein[J]. J Cell Physiol, 1993, 157(1):33–40. doi: 10.1002/jcp.1041570105.
- [37] Yang Y, Li F, Luo X, et al. Identification of LCN1 as a Potential

- Biomarker for Breast Cancer by Bioinformatic Analysis[J]. DNA Cell Biol, 2019, 38(10):1088–1099. doi: 10.1089/dna.2019.4843.
- [38] Mullen EM, Gu P, Cooney AJ. Nuclear Receptors in Regulation of Mouse ES Cell Pluripotency and Differentiation[J]. PPAR Res, 2007, 2007:61563. doi: 10.1155/2007/61563.
- [39] Zhao H, Li Z, Cooney AJ, et al. Orphan nuclear receptor function in the ovary[J]. Front Biosci, 2007, 12:3398–3405. doi: 10.2741/2321.
- [40] Cheng G, Wang S, Li X, et al. Positive expression of NR6A1/CT150 as a predictor of biochemical recurrence-free survival in prostate cancer patients[J]. Oncotarget, 2016, 8(38):64427–64439. doi: 10.18632/oncotarget.11749.
- [41] Jüttner R, Montag D, Craveiro RB, et al. Impaired presynaptic function and elimination of synapses at premature stages during

- postnatal development of the cerebellum in the absence of CALEB (CSPG5/neuroglycan C)[J]. Eur J Neurosci, 2013, 38(9):3270–3280. doi: 10.1111/ejn.12313.

(本文编辑 宋涛)

本文引用格式: 彭颖, 龚光辉, 李景和, 等. 基于免疫相关基因标签的原发性肝癌预后评分系统的建立[J]. 中国普通外科杂志, 2020, 29(2):179–189. doi:10.7659/j.issn.1005–6947.2020.02.009

Cite this article as: Peng Y, Gong GH, Li JH, et al. Establishment of a prognostic risk score model of hepatocellular carcinoma based on an immune-related gene signature[J]. Chin J Gen Surg, 2020, 29(2):179–189. doi:10.7659/j.issn.1005–6947.2020.02.009

本刊常用词汇英文缩写表

C- 反应蛋白	CRP	甲型肝炎病毒	HAV	心电图	ECG
Toll 样受体	TLRs	碱性成纤维细胞转化生长因子	bFGF	心脏监护病房	CCU
氨基末端激酶	JNK	聚合酶链反应	PCR	血管紧张素 II	AngII
白细胞	WBC	抗生物素蛋白 - 生物素酶复合物法	ABC 法	血管内皮生长因子	VEGF
白细胞介素	IL	辣根过氧化物酶	HRP	血管性血友病因子	vWF
半数抑制浓度	IC ₅₀	链霉抗生物素蛋白 - 生物素酶复合物法	SABC 法	血红蛋白	Hb
变异系数	CV	磷酸盐缓冲液	PBS	血肌酐	SCr
标记的链霉抗生物素蛋白 - 生物素法	SP 法	绿色荧光蛋白	GFP	血尿素氮	BUN
表皮生长因子	EGF	酶联免疫吸附测定	ELISA	血小板	PLT
丙氨酸氨基转移酶	ALT	美国食品药品监督管理局	FDA	血压	BP
丙二醛	MDA	脑电图	EEG	血氧饱和度	SO ₂
丙型肝炎病毒	HCV	内毒素 / 脂多糖	LPS	烟酰胺腺嘌呤二核苷酸	NADPH
超氧化物歧化酶	SOD	内皮型一氧化氮合酶	eNOS	严重急性呼吸综合征	SARS
磁共振成像	MRI	内生肌酐清除率	CCr	一氧化氮	NO
极低密度脂蛋白胆固醇	VLDL-C	尿素氮	BUN	一氧化氮合酶	NOS
低密度脂蛋白胆固醇	LDL-C	凝血酶时间	TT	乙二胺四乙酸	EDTA
动脉血二氧化碳分压	PaCO ₂	凝血酶原时间	PT	乙酰胆碱	ACh
动脉血氧分压	PaO ₂	牛血清白蛋白	BSA	乙型肝炎病毒	HBV
二甲亚砜	DMSO	热休克蛋白	HSP	乙型肝炎病毒 e 抗体	HBeAb
反转录 - 聚合酶链反应	RT-PCR	人类免疫缺陷病毒	HIV	乙型肝炎病毒 e 抗原	HBeAg
辅助性 T 细胞	Th	人绒毛膜促性腺激素	HCG	乙型肝炎病毒表面抗体	HBsAb
肝细胞生长因子	HGF	三磷酸腺苷	ATP	乙型肝炎病毒表面抗原	HBsAg
干扰素	IFN	三酰甘油	TG	乙型肝炎病毒核心抗体	HBeAb
高密度脂蛋白胆固醇	HDL-C	生理氯化钠溶液	NS	乙型肝炎病毒核心抗原	HBeAg
谷胱甘肽	GSH	世界卫生组织	WHO	异硫氰酸荧光素	FLTC
固相 pH 梯度	IPG	双蒸水	ddH ₂ O	诱导型一氧化氮合酶	iNOS
核糖核酸	RNA	丝裂原活化蛋白激酶	MAPK	原位末端标记法	TUNEL
核因子 - κ B	NF-κ B	四甲基偶氮唑盐微量酶反应	MTT	杂合性缺失	LOH
红细胞	RBC	苏木精 - 伊红染色	HE	增强化学发光法	ECL
红细胞沉降率	ESR	胎牛血清	FBS	肿瘤坏死因子	TNF
环氧酶 -2	COX-2	体质量指数	BMI	重症监护病房	ICU
活化部分凝血活酶时间	APTT	天门冬氨酸氨基转移酶	AST	转化生长因子	TGF
活性氧	ROS	脱氧核糖核酸	DNA	自然杀伤细胞	NK 细胞
获得性免疫缺陷综合征	AIDS	细胞间黏附分子	ICAM	直接胆红素	DBIL
肌酐	Cr	细胞外基质	ECM	总胆固醇	TC
基质金属蛋白酶	MMP	细胞外调节蛋白激酶	ERK	总胆红素	TBIL
计算机 X 线断层照相技术	CT	纤连蛋白	FN		